Differentiable Selections of Set-Valued Mappings With Application in Stochastic Programming *

Darinka Dentcheva
Humboldt University Berlin, Institute of Mathematics
Unter den Linden 6 ,10099 Berlin, Germany
dary@mathematik.hu-berlin.de

Abstract

We consider set-valued mappings defined on a linear normed space with convex closed images in $I\!R^n$. Our aim is to construct selections which are (Hadamard) directionally differentiable using some approximation of the multifunction. The constructions suggested assume existence of a cone approximation given by a certain "derivative" of the mapping. The first one makes use of the properties of Steiner points. The notion of Steiner center is generalized for a class of unbounded sets, which include the polyhedral sets. The second construction defines a continuous selection through a given point of the graph of the multifunction and being Hadamard directionally differentiable at that point with derivatives belonging to the corresponding "derivative" of the multifunction. Both constructions lead to a directionally differentiable Castaing representation of measurable multifunctions with the required differentiability properties. The results are applied to obtain statements about the asymptotic behaviour of measurable selections of random sets via the delta-approach. Particularly, random sets of this kind build the solutions of two-stage stochastic programs.

Keywords: differentiable set-valued mapping, Steiner center, selections, Castaing representation, delta-theorem, two-stage stochastic programs

1 Introduction

Analysis of the behaviour of set-valued mappings (multifunctions) gives rise to a natural question: "Can we find selections of the mapping under consideration inheriting their regularity properties?"
Let $F : X \rightrightarrows I\!R^n$ be a multifunction acting from a linear normed space X in $I\!R^n$. A

*Supported by the Deutsche Forschungsgemeinschaft

mapping $f : X \to I\!R^n$ is a *selection* of F if $f(x) \in F(x)$ for all $x \in X$.
By the axiom of choice, such a mapping exists whenever F has non-empty images. One is usually interested in finding a selection with some additional properties, like measurability, continuity, differentiability etc., motivated by the specifics of the particular investigation.
This paper is devoted to the question how differentiability properties of a multifunction are inherited by its selections.
We would like to review some of the results on selections with remarkable properties. Our brief summary is not an attempt to comment all the interesting results published on this topic, we just want to mention some of them in order to provide an initial inside into the subject. One of the most celebrated results on the existence of continuous selections is the following theorem of Michael [25]. If X is a paracompact space and F is a lower semi-continuous mapping with closed convex images, then F admits a continuous selection. The key step in the proof of this theorem is the construction of a continuous ε-approximate selection. In [14], Deutsch and Kenderov provide a topological characterization of those mappings that admit continuous ε-approximate selection. Existence of Lipschitz-continuous selections has been shown in [5] by Aubin and Cellina.
Fundamental results of existence of measurable selections are contained in [10, 11, 23]. A nice overview of the basic facts how selections inherit measurability, Lipschitz-continuity etc. is given in [6]. The reader can find there also a presentation of some special selections and their properties which are widely studied in the literature.
The stronger is the regularity one requires for the selections, the more are the technical difficulties one is facing. Asking for differentiability is motivated both by theoretical as well as by practical reasons as we shall see. It is natural to try to relate some differentiability property of a set-valued mapping with the differentiability of some selection of it. When dealing with some mapping expressing solutions of optimization problems, one usually cannot control all solutions. One is interested in computing or in having some information about a selection of it.
Several concepts of differentiability of set-valued mappings have been developed in the literature. The aim is to gain information of use about the mapping in connection with optimization and other allied subjects. Some pioneering geometrical concepts are presented in [6], which correspond to looking at certain tanget cone to the graph of the multifunction. The contingent cone and the Clarke-tangent cone are considered in [3, 4, 5, 6] and an intermediate type of tangent cone is first treated by Urcescu [41] and Frankowska [17, 18].
A stronger differentiability notion, called proto-differentiability, has-been introduced by Rockafellar [31]. It is related to interesting phenomena in optimization as it is shown in [28, 29, 31]. To mention briefly some of them, proto-differentiability is commonly present for the mapping, expressing feasibility in optimization problems, when certain constraint qualification is satisfied. It is also a characteristic property of the subgradient mapping of a convex function, whenever this function is twice epi-differentiable.
Another attractive concept of differentiability of multifunctions is the semi-differentiability. This concept goes back to Penot [26] and has been developed further by Rockafellar [31]. Semi-differentiability plays an important role in the so-called delta-approach,

which provides information about the asymptotic behaviour of stochastic processes. Delta-theorems establish central limit formulas for random elements, respectively random sets in case of measurable multifunctions. Mappings containing optimal solutions of stochastic programs are, particularly, of this kind.
Our intention is to establish some relations between the differentiability of a set-valued mapping in the sense above and the existence of selections with corresponding differentiability properties. On the way to this end, we extend the notion of a Steiner center for a class of unbounded sets and obtain another result on the existence of Lipschitz-continuous selections.
The existence of a differentiable selection has been treated in [19, 15]. Gautier and Morchadi consider the so-called Steiner selection and its remarkable properties in [19]. They assume to have an affine approximation of the set-valued mapping and show that the Steiner selection is differentiable then. This relation is similar to those we are aiming at. The problem is that the existence of an affine approximation might be very hard to establish. In [15], Dommisch establishes two results on the topic. The existence of a smooth partition of unity in $\mathbb{R}^n$ is exploited there to prove the existence of a smooth selection under some condition we will comment at the end of section 4.

We follow both approaches in our investigations. Both, Steiner selections as well as selections constructed via the partition of unity will be considered. Both constructions preserve measurability and lead to a Cataing representation with differentiability properties of the measurable multifunction under consideration. It has nontrivial consequences for the convergence in distribution of selections of random sets. This problem is widely studied in the literature (see [1, 2] and the references therein). The paper [1] of Artstein provides some fundamental results on this subject. We will discuss the matter later on. In this place we only want to recall that the distribution of a random set does not determine the distributions of its selections.
An application in two-stage stochastic programming is given.

In section 2, we formulate the necessary notions and recall some preliminary facts.
In section 3, we propose a slight generalization of the notion of a Steiner center. This leads to a result about the existence of Lipschitz-continuous selections of set-valued mappings with unbounded images, which are not necessarily Lipschitzian themselves. Further, we show how the Steiner selection inherits the differentiability of the multifunction. Under certain assumptions directional, proto- and semi-differentiability lead to a directional, respectively Hadamard-directional differentiability of the Steiner selection.
Section 4 contains a construction of a Hadamard directionally differentiable selection of a mapping, which is semi-differentiable at some points. It exploits the existence of a continuous partition of unity.
In section 5, measurability properties, Castaing representation theorems and convergence in distribution of selection of random sets are discussed.
In section 6, we demonstrate how the previous results, applied to solution sets mappings of two-stage stochastic programs, might provide information about the asymptotic distributions of selections of solution sets.

A last remark is due. It is possible to prove some of the presented results also for mappings valued in infinite dimensional spaces. This topic exceeds the scope of this paper and will be investigated elsewhere.

2 Differentiability of multifunctions

Throughout the whole paper we will consider set-valued mappings $F : X \rightrightarrows I\!R^n$ defined on a linear normed space X and having values in $I\!R^n$.
We denote the graph of F by $graphF$.
A general assumption needed is that F has non-empty images, i.e., $F(x) \neq \emptyset$ for all $x \in X$.
The following notions of differentiability of set-valued mappings will be used.

Definition 2.1 *The mapping $F : X \rightrightarrows I\!R^n$ is called directionally differentiable at a point $(\bar{x}, \bar{y}) \in graphF$ in direction $h \in X$, if the limit*

$$F'(\bar{x}, \bar{y}; h) = \lim_{t \downarrow 0} \frac{F(\bar{x} + th) - \bar{y}}{t}$$

exists in the sense of Kuratowski-Painlevé convergence.

Recall that a sequence of closed sets $\{A_n\}, A_n \subseteq I\!R^n$ converges to some $A \subseteq I\!R^n$ in the sense of Kuratowski-Painlevé (written $A_n \to A$) if and only if the sequence of the distance functions converges pointwise (cf. [6]), i.e.,

$$A_n \to A \quad \text{if and only if} \quad d(y, A_n) \to d(y, A) \text{ for all } y \in I\!R^n,$$

where

$$d(y, A) = \inf\{\|z - y\| \ : \ z \in A\}.$$

$$\liminf_{n \to \infty} A_n = \{z \ : \ \limsup_{n \to \infty} d(z, A_n) = 0\}$$

$$\limsup_{n \to \infty} A_n = \{z \ : \ \liminf_{n \to \infty} d(z, A_n) = 0\}$$

and $A_n \to A$ means $\liminf_{n \to \infty} A_n = \limsup_{n \to \infty} A_n = A$.
Let $(x, y) \in graphF$ be a given point. We use the notation

$$D_t F(h) = t^{-1}[F(x + th) - y] \quad \text{for } t > 0.$$

Definition 2.2 *The mapping $F : X \rightrightarrows I\!R^n$ is said to be proto-differentiable at a point $(x, y) \in graphF$, if there is a set-valued mapping $DF : X \rightrightarrows I\!R^n$ such that $D_t F$ converges in graph to DF while $t \downarrow 0$, i.e.,*

$$graphDF = \lim_{t \downarrow 0} graphD_t F \quad in \quad X \times I\!R^n.$$

We denote the proto-derivative (the mapping DF) by $DF(x,y)$ to indicate the dependence on the chosen point.

Definition 2.3 *The mapping $F : X \rightrightarrows \mathbb{R}^n$ is called semi-differentiable at a point $(x,y) \in graph F$, if the limit*

$$DF(x,y;h_0) = \lim_{t \downarrow 0, h \to h_0} \frac{F(x+th)-y}{t}$$

exists for all $h_0 \in X$.

The different differentiability concepts are compared in [7, 31]. In [31], it has been established that, if a mapping is directionally respectively proto-differentiable and pseudo-Lipschitzian, then it is semi-differentiable, too. The semi-differentiability has the advantage to generate a derivative that forms a continuous multifunction with respect to the direction (cf. [7]).
All derivatives above build some cone-approximation of the graph of the multifunction. The latter one is continuous.
Continuous tangential approximations of set-valued mappings are considered also in [36, 37]. It has been shown in [7] that such tangential approximations, if they exist, coincide with the semi-derivatives.

3 Steiner selections

3.1 A Steiner center for asymptotically affine sets

In this section, we introduce the notions of an asymptotically affine set and a Steiner center for such kind of sets.
Let $C \subseteq \mathbb{R}^n$ be a closed convex set. We consider the representation of C as a Minkowski sum of its recession cone and a convex hull of a set, which possesses extreme points. Recall that the recession cone $K(C)$ and the recession subspace $S(C)$ are defined by

$$K(C) = \{y \in \mathbb{R}^n \ : \ x + \lambda y \in C \ \ \forall x \in C \ \forall \lambda \in \mathbb{R}_+\}$$

$$S(C) = \{z \in \mathbb{R}^n \ : \ x + \lambda z \in C \ \ \forall x \in C \ \forall \lambda \in \mathbb{R}\}$$

Further we denote by $T(C)$ the complement subspace of $S(C)$ and by $\hat{C}$ the set of extreme points of C.
It is a well-known fact from convex analysis that the set $C_T := C \cap T(C)$ possesses extreme points, i.e., $\hat{C}_T \neq \emptyset$ and the set C can be represented as follows:

$$C = co\hat{C}_T + K(C), \tag{1}$$

where $co\hat{C}_T$ denotes the convex hull of $\hat{C}_T$.
Using the representation we impose restrictions on $\hat{C}_T$ to introduce the asymptotically affine sets.

Definition 3.1 *We call a closed convex set $C \subseteq \mathbb{R}^n$ asymptotically affine if the corresponding set $\hat{C}_T$ is bounded.*

We denote the set $co(\hat{C}_T)$ by C_S. It is evident that C_S is a compact set.

Example 3.2 *All compact sets are asymptotically affine, obviously.*

Example 3.3 *All polyhedral sets are asymptotically affine by Theorem 19.1 in [32]*

Example 3.4 *All "cylindric" sets of the form $B(y,\alpha) \times \mathbb{R}$ for some $y \in \mathbb{R}^n$ or, more generally, all sets of the form $C \times \mathbb{R}^k$, $0 \leq k < n$, where C is a compact subset of $\mathbb{R}^{n-k}$.*

We proceed with the definition and properties of the Steiner center.

Definition 3.5 *Given a compact convex set $C \subseteq \mathbb{R}^n$, the Steiner center $st(C)$ of C is defined in the following way:*

$$st(C) = \int_{\Sigma} np\sigma(p,C)\alpha(dp),$$

where
Σ denotes the unit sphere in $\mathbb{R}^n$,
α is a normalized isometry-invariant measure on Σ,
$\sigma(\cdot, C)$ is the support function of C.

Recall that the support function $\sigma(\cdot, C) : \mathbb{R}^n \to \mathbb{R}$ of a closed convex set $C \subseteq \mathbb{R}^n$ is defined by

$$\sigma(p,C) = \sup_{y \in C} \langle p, y \rangle.$$

This point was first introduced by Steiner [40] in 1840 for a C^2-convex plane curve as the barycenter of the curvature measure. The definition given above was introduced by Shepard [39]. As a matter of fact, we can use any probability measure α on Σ in the defining formula and the properties of the Steiner center used for our purposes still remain true. We recall these properties as well as the properties of the support function which determine to great extend those mentioned first.

- $\sigma(\cdot, C)$ is finite for all $p \in \mathbb{R}^n$ and for a compact convex set C,
- $\sigma(p, \lambda C) = \lambda\sigma(p,C)$ for all $p \in \mathbb{R}^n$, for all $\lambda \geq 0$
- $\sigma(p, C + C_1) = \sigma(p,C) + \sigma(p,C_1)$, where C and C_1 are compact convex sets.
- $\sigma(p,C) \leq \sigma(p,C_1)$ for $C \subseteq C_1$ and C and C_1 are compact convex sets.

Throughout the paper we denote the Hausdorff distance between two sets $A, B \subseteq I\!R^n$ by

$$d_H(a,b) = \max\{e(A,B), e(B,A)\},$$

where

$$e(A,B) = \sup_{y \in A} d(y,B)$$

Basic properties of the Steiner center are the following ones:

1. $st(C) \in C$ for all compact convex sets $C \subseteq I\!R^n$.
2. $st(\lambda A + \mu B) = \lambda st(A) + \mu st(B)$ for any real numbers λ and μ and any compact convex sets A and B.
3. The mapping $C \mapsto st(C)$ defined for all nonempty convex compact sets is Hausdorff-Lipschitz:
$$\|st(A) - st(B)\| \leq L d_H(A,B).$$
The best value for the Lipschitz-constant L has been proved (cf.[34]) to be the inverse of the Wallis integral, i.e.,
$$L = \frac{2}{\pi^{1/2}} \Gamma(\frac{n+2}{2}) \Gamma(\frac{n+1}{2})^{-1}.$$

Definition 3.6 *A Steiner senter $st(C)$ of an asymptotically affine set C is the Steiner center of the corresponding set C_S, i.e. $st(C) = st(C_S)$.*

Note that $st(C) \in C_S \subseteq C$. Moreover, if C is a compact set, the $C_S \equiv C$ and thus the Steiner center coincides with the point from the first definition. The property 2., i.e. the affine character of the mapping $C \mapsto st(C)$, remains valid for asymptotically affine sets, too. The property 3. has to be modified. The Hausdorff distance between two unbounded sets A and B could be finite in the special situation, when $K(A) = K(B)$. This is, however, not necessary for obtaining a Lipschitz behaviour of the Steiner centers.

Proposition 3.7 *The mapping $A \mapsto st(A)$ defined for all nonempty asymptotically affine sets has Lipschitz behaviour in the following sense:*

$$\|st(A) - st(B)\| \leq L d_H(A_S, B_S).$$

with the same Lipschitz-constant as in 3.

Proof: The statement is an immediate consequence of the definitions and property 3. □

The observation in this proposition is very easy, but it has nontrivial consequences for the existence of Lipschitz selections of set-valued mappings.
Let us consider a multifunction $F : X \rightrightarrows I\!R^n$, where X is a linear normed space. In the remainder of this section we suppose that for all $x \in X$, $F(x)$ are non-empty asymptotically affine sets.

Definition 3.8 *The function $f : X \to I\!R^n$, defined by*

$$f(x) = st([F(x)]_S)$$

is said to be a Steiner selection of F.

If, in addition, F is a locally Lipschitzian mapping, i.e., for any $\bar{x} \in X$ there exist a constant L and a neighbourhood U of $\bar{x}$ such that

$$d_H(F(x_1), F(x_2)) \leq L\|x_1 - x_2\| \text{ for all } x_1, x_2 \in U,$$

then the Steiner selection is locally Lipschitzian by the result of Aubin and Cellina [5]. Results about existence of Lipschitz-continuous selections are given in [5, 15, 6], including the case of $F(X)$ being unbounded sets. The results in [15] suppose X to be a compact metric space, and among other assumptions, one of the following two is required:

- $F^{-1}(y) = \{x \in X : y \in F(x)\}$ is open in X for all $y \in I\!R^n$.
- int$F(x) \neq \emptyset$ for all $x \in X$.

where int$A$ denotes the interior of A. Among general results, we are aiming at some application of them in optimization. The graph of solution sets mappings of perturbed (parametric) optimization problems is closed under relatively mild assumptions and solution sets have non-empty interior very seldom.
We will call a multifunction F_s with the property that $graphF_s \subseteq graphF$ *a subfunction* of F. It is clear that if there exists a locally Lipschitzian subfunction of F with compact values then F_s and, thus, also F has a locally Lipschitz-continuous selection. We consider the subfunction F_s defined by

$$F_s(x) = co[\widehat{F(x)}]_T \qquad \text{(cf. 1)} \tag{2}$$

F_s is well-defined.
The example below shows that none of the requirements for the results cited above is satisfied, i.e.,
int$F(x) = \emptyset$ for all $x \in X$, $F^{-1}(y)$ is closed for all $y \in I\!R^n$ and $d_H(F(x), F(y)) = \infty \ \forall x \neq y$. But the corresponding subfunction F_s in the example is locally Lipschitzian and, consequently, the Steiner selection of F is Lipschitz-continuous by the Proposition.

Example 3.9 *$F : I\!R \rightrightarrows I\!R^3$ is defined by setting*

$$F(x) = \{(y_1, y_2, y_3) : \begin{array}{l} sinx.y_1 + cosx.y_2 \geq 0, \\ sinx.y_1 - cosx.y_2 \leq 0, \\ y_3 = x \ \} \end{array}$$

It is easy to see that the Steiner selection of F is $f(x) = (0, 0, x)$.

3.2 Directional differentiability of the Steiner selection

In this section, we give sufficient conditions for the (Hadamard-) directional differentiability of the Steiner selection. Our basic assumption is that the multifunction $F : X \rightrightarrows I\!R^n$ under consideration has non-empty asymptotically affine values. At the end of this section we will make a comparison between the results known from the literature and those stated here.
First, we prove a statement that will be applied in the proof of the main result, but might be a useful and interesting observation independently.

Lemma 3.10 *Let $F : X \rightrightarrows I\!R^n$ be compact valued and locally upper Lipschitzian at $\bar{x}$, i.e., there exists a constant L and a neighbourhood U of $\bar{x}$:*

$$d_H(F(x), F(\bar{x})) \leq L\|x - \bar{x}\| \quad \forall x \in U$$

Let $\sigma(p, F(\bar{x}))$ be the support function of $F(\bar{x}), \|p\| = 1$ and $\sigma(p, F(\bar{x})) = \langle p, \bar{y}\rangle$ for some $\bar{y} \in F(\bar{x})$. Suppose that F is semi-differentiable at the point $(\bar{x}, \bar{y})$ with the semi-derivative $DF(\bar{x}, \bar{y}; h)$ in direction h.
Then $\sigma(p, DF(\bar{x}, \bar{y}; h))$ is finite and bounded by $L\|h\|$.

Proof:
Claim 1: $\langle p, z\rangle \leq 0$ for all $z \in T_{F(\bar{x})}(\bar{y})$, where $T_{F(\bar{x})}(\bar{y})$ is the tangent cone of $F(\bar{x})$ at the point $\bar{y}$.
Recall that $T_{F(\bar{x})}(\bar{y}) = \lim_{t \downarrow 0} \frac{1}{t}(F(\bar{x}) - \bar{y})$, where the limit is understood in the sense of Kuratowski-Painlevé. Assume the opposite to be true, i.e.

$$\langle p, z\rangle = c > 0 \text{ for some } \quad z \in T_{F(\bar{x})}(\bar{y}).$$

Then there exist sequences $\{z_i\}$, $\{t_i\}$, $i = 1, 2, \ldots$ such that $t_i \downarrow 0$, $z_i \to z$ and $y_i = \bar{y} + t_i z_i \in F(\bar{x})$. By $c > 0$, there exists an i_0 such that $\langle p, z_i\rangle > 0 \ \forall i \geq i_0$.
Consequently, $\langle p, y_i\rangle = \langle p, \bar{y}\rangle + t_i\langle p, z_i\rangle > \langle p, \bar{y}\rangle$ for i large enough, which is a contradiction to the definition of a support function.
Claim 2: $e(DF(\bar{x}, \bar{y}; h), T_{F(\bar{x})}(\bar{y})) \leq L\|h\|$.
Let $z \in DF(\bar{x}, \bar{y}; h)$. By definition there exists sequences $\{h_i\}$, $\{t_i\}, \{y_i\}$, $\{z_i\}$ $i = 1, 2, \ldots$ such that $t_i \downarrow 0$, $h_i \to h$, $z_i \to z$ and $y_i = \bar{y} + t_i z_i \in F(\bar{x} + t_i h_i)$.
We denote the metric projection of the point y on the set A by $pr_A y$. Let us, in particular, denote $\bar{y}_i = pr_{F(\bar{x})} y_i$. It is well-known that $\|pr_A y - \hat{y}\| \leq \|y - \hat{y}\| \ \forall \hat{y} \in A$, and hence

$$\|\bar{y}_i - \bar{y}\| \leq \|y_i - \bar{y}\| \leq t_i\|z_i\|.$$

This implies that $\|(\bar{y}_i - \bar{y})/t_i\|$ is a bounded sequence and, consequently, it has a convergent subsequence:

$$(\bar{y}_{i_k} - \bar{y})/t_{i_k} \to \bar{z}.$$

Obviously, $\bar{z} \in T_{F(\bar{x})}(\bar{y})$. The following sequence of inequalities holds true:

$$\|\bar{z} - z\| \quad = \quad lim_{k\to\infty}\|\frac{\bar{y}_{i_k} - \bar{y}}{t_{i_k}} - \frac{y_{i_k} - \bar{y}}{t_{i_k}}\|$$

$$
\begin{aligned}
&= lim_{k\to\infty}\|\frac{y_{i_k}-\bar{y_{i_k}}}{t_{i_k}}\| \\
&\leq lim_{k\to\infty}\frac{1}{t_{i_k}}d_H(F(\bar{x}+t_{i_k}h_{i_k}),F(\bar{x})) \\
&\leq lim_{k\to\infty}L\|h_{i_k}\| = L\|h\|
\end{aligned}
$$

The two assertions imply the desired estimate for $\sigma(p, DF(\bar{x},\bar{y};h))$:

$$
\begin{aligned}
\sigma(p,DF(\bar{x},\bar{y};h)) &= \max\{\langle p,z\rangle :\ z\in DF(\bar{x},\bar{y};h)\} \\
&= \max\{\langle p,z-\bar{z}\rangle+\langle p,\bar{z}\rangle :\ z\in DF(\bar{x},\bar{y};h),\ \bar{z}=pr_T z\in T_{F(\bar{x})}(\bar{y})\} \\
&\leq \max_z\{\|p\|\|z-\bar{z}\|+\langle p,\bar{z}\rangle\} \\
&\leq L\|h\|
\end{aligned}
$$

□

Recall that $DF(\bar{x},\bar{y};0) = T_{F(\bar{x})}(\bar{y})$ and $DF(\bar{x},\bar{y};\cdot)$ is positive homogeneous. Observe that $DF(\bar{x},\bar{y};h)\neq\emptyset$ for all h, by the Lipschitz-property and the assumption that $F(x)\neq\emptyset$ for all $x\in X$.

Theorem 3.11 *Given a multifunction $F: X\rightrightarrows\mathbb{R}^n$, suppose that*

- *the corresponding subfunction F_s (cf. 2) is locally upper Lipschitzian at all points $x\in X$,*
- *F_s is semi-differentiable at all points $(x,y)\in graphF$ such that $y\in bdF(x)$, where bd stands for the boundary of $F(x)$.*

Then the Steiner selection f of F is Hadamard-directionally differentiable at all points $x\in X$. Moreover, the directional derivative of F is given by the following formula:

$$
f'(x;h) = \int_\Sigma np\sigma(p,DF_s(x,y_p;h)\alpha(dp), \tag{3}
$$

where $y_p\in F_s(x):\quad \sigma(p,F_s(x)) = \langle p,y_p\rangle$.

Proof: Given a direction h_0 and $p\in\Sigma$, we consider

$$
\lim_{t\downarrow 0,h\to h_0}\frac{f(x+th)-f(x)}{t} = \lim_{t\downarrow 0,h\to h_0}\int_\Sigma np\frac{\sigma(p,F_s(x+th))-\sigma(p,F_s(x))}{t}\alpha(dp) \tag{4}
$$

We denote by $\bar{y}$ and $\bar{y}_t$ the points from $F_s(x)$ and $F_s(x+th)$, respectively, such that $\sigma(p,F_s(x)) = \langle p,\bar{y}\rangle$ and $\sigma(p,F_s(x+th)) = \langle p,\bar{y}_t\rangle$. It is clear that $\bar{y}\in bdF(x)$. Consider $DF_s(x,\bar{y};h_0)$. By the virtue of the previous lemma $\langle p,z\rangle$ is bounded for $z\in DF_s(x,\bar{y};h_0)$. Let us denote by $\bar{z}$ the maximum point,i.e.,

$$
\sigma(p,DF_s(x,\bar{y};h_0)) = \langle p,\bar{z}\rangle.
$$

The following estimation holds due to the first assumption and the properties of the support function:

$$\begin{aligned} & t^{-1}[\sigma(p, F_s(x+th)) - \sigma(p, F_s(x))] \\ = \;& t^{-1}\langle p, \bar{y}_t - \bar{y}\rangle \\ \leq \;& L\|h\| \end{aligned}$$

Consequently, we can exchange the limit and the integral in 4 by Lebesque Theorem. Thus, f has the same differentiability properties as the support function. We consider $\limsup t^{-1}[\sigma(p, F_s(x+th)) - \sigma(p, F_s(x))]$. The sequence $t^{-1}(\bar{y}_t - \bar{y})$ is bounded and, hence it has a converging subsequence $t_i^{-1}(\bar{y}_{t_i} - \bar{y})$, for which the lim sup is attained:

$$t_i^{-1}(\bar{y}_{t_i} - \bar{y}) \to z \text{ while } \; t_i \downarrow 0, h_i \to h_0$$

Consequently, $z \in DF(x, \bar{y}; h)$. Thus

$$\begin{aligned} & \limsup_{t\downarrow 0, h\to h_0} t^{-1}[\sigma(p, F_s(x+th)) - \sigma(p, F_s(x))] \qquad (5) \\ = \;& \lim_{i\to\infty} t_i^{-1}\langle p, \bar{y}_{t_i} - \bar{y}\rangle \\ = \;& \langle p, z\rangle \leq \langle p, \bar{z}\rangle \end{aligned}$$

On the other hand we consider the points $y_t \in F_s(x+th)$, $y_t = \bar{y} + tz_t$ with $z_t \to \bar{z}$ for $t \downarrow 0, h \to h_0$, which exist by the definition of a semi-derivative.

$$\begin{aligned} & \liminf_{t\downarrow 0, h\to h_0} \frac{1}{t}[\sigma(p, F_s(x+th)) - \sigma(p, F_s(x))] \qquad (6) \\ = \;& \liminf_{t\downarrow 0, h\to h_0} \frac{1}{t}(\max_{F_s(x+th)} \langle p, y\rangle - \langle p, \bar{y}\rangle) \\ \geq \;& \liminf_{t\downarrow 0, h\to h_0} \frac{1}{t}(\langle p, y_t\rangle - \langle p, \bar{y}\rangle) \\ = \;& \liminf_{t\downarrow 0, h\to h_0} \langle p, z_t\rangle = \langle p, \bar{z}\rangle \end{aligned}$$

The inequalities 5 and 6 imply the assertion. □

The notion of semi-differentiability and upper Lipschitz-continuity have a local character, so that the question of a "local version" of the theorem above is of interest. Moreover, to verify these properties for particular mappings might be a difficult task. A similar question arises also if F_s has empty images at some points. Such a local version is contained in the next corollary, the proof of which follows from the proof of the previous theorem.

Corollary 3.12 *Let $F : X \rightrightarrows \mathbb{R}^n$ and a point $\bar{x} \in X$ be given. Suppose the corresponding subfunction F_s to be Lipschitzian (locally upper Lipschitzian) at the point $\bar{x}$ and semi-differentiable at all points $(\bar{x}, y)$, $y \in bdF_s(\bar{x})$.*
Then the Steiner selection f of F is a Lipschitz-continuous (locally upper Lipschitz-continuous) mapping. f is Hadamard-directionally differentiable at $\bar{x}$ and the directional derivative satisfies the formula 3.

Now, we would like to formulate a statement relating the directional and proto- differentiability of a set-valued mapping with the directional differentiability of the Steiner selection.

Corollary 3.13 *Given a multifunction $F : X \rightrightarrows \mathbb{R}^n$, suppose the corresponding subfunction F_s to be locally upper Lipschitzian (Lipschitzian) and directionally or proto-differentiable at all points $(x, y) \in graph F_s : \ y \in bd F_s(\bar{x})$.*
Then the Steiner selection f of F is (Hadamard)-directionally differentiable and the directional derivative satisfies the formula 3.
If the assumptions hold only for a fixed point $\bar{x}$, then the Steiner selection is (Lipschitz-) continuous and (Hadamard)-directionally differentiable at $\bar{x}$.

Proof: The proof is the same as that of Theorem 3.11. We have to take into account that directional differentiability together with Lipschitz-continuity imply semi-differentiability, as it has been proven in [31]. In case of upper Lipschitz-continuity we follow the same proof considering all limits for fixed direction h and obtain the directional differentiability of the selection. □

At the end of this section we compare the results above with those presented in [19]. It has been shown there that the Steiner selection is differentiable if $F(x)$ are non-empty compact convex sets and F has an affine approximation A such that $A(x)$ are non-empty compact convex sets,too. The semi-derivatives or the directional derivatives build a cone-approximation of F and it cannot be expected that some of them is an affine approximation. Moreover, these derivatives are always unbounded in case $F(x)$ is not a singleton. Semi- and directional derivatives are established approximations, studied and used in the literature (cf. [7, 12, 31, 37], etc.) for different purposes. This is the reason we gave our preference to them.

4 Existence of directionally differentiable selections via partition of unity

This section is devoted to another approach for establishing a relation between the differentiability properties of a set-valued mapping and its selections. This is a general tool of analysis and very often used while constructing selections. We exploit the existence of a continuous partition of unity, subordinated to a given open covering of a linear normed space.
We recall that, given a linear normed space X and an open covering $X = \cup_{\alpha\in\Lambda} V_\alpha$ of it, a continuous partition of unity subordinated to $\{V_\alpha\}_\Lambda$ is an union $\{g_\alpha\}_\Lambda$ of continuous functions g_α such that:

- $0 \leq g_\alpha(x) \leq 1$ for all $x \in X$,
- $\cup_{\alpha\in\Lambda}(supp g_\alpha) = X$,
- for any $\alpha \in \Lambda$ there exists a finite number of $\beta \in \Lambda$ such that $supp g_\alpha \cap supp g_\beta \neq \emptyset$,

- for any $\alpha \in \Lambda$ it holds $supp g_\alpha \subseteq V_\alpha$.

Of course, the existence of such a partition is known for a wider class of spaces, but this formulation corresponds to our setting.
We first formulate two lemmas.

Lemma 4.1 *Let $F : X \rightrightarrows \mathbb{R}^n$ be semi-differentiable at $(\bar{x}, \bar{y})$. Then the semi-derivative is a continuous function of h in the sense that*

$$\liminf_{h \to h_0} DF(\bar{x}, \bar{y}; h) = \limsup_{h \to h_0} DF(\bar{x}, \bar{y}; h) = DF(\bar{x}, \bar{y}; h_0).$$

Proof: The assertion has been proved in [7] for the case where X is a finite-dimensional space. The proof works also for linear normed spaces and is omitted here. □

We denote the metric projection of $y \in \mathbb{R}^n$ onto a closed convex set A by $pr_A y$.

Lemma 4.2 *Given a multifunction $F : X \rightrightarrows \mathbb{R}^n$ and a point $\bar{x} \in X$ such that $F(x) \underset{x \to \bar{x}}{\longrightarrow} F(\bar{x})$, suppose that the single-valued mapping $y : X \to \mathbb{R}^n$ is continuous at $\bar{x}$. Then the metric projection $f(x) = pr_{F(x)} y(x)$ is continuous at $\bar{x}$.*

Proof: We choose some positive number ρ_0 and set

$$\bar{\rho} = \max\{1/6, \|y(\bar{x})\| + 2d(y(\bar{x}), F(\bar{x})) + \rho_0\}$$

Following the monograph [33], we define

$$\hat{d}_\rho(C, D) = \inf\{\eta \geq 0 : C \cap \rho B \subseteq D + \eta B, D \cap \rho B \subseteq C + \eta B\}$$

as the ρ-distance between the sets C and D. Here B denotes the closed unit ball in $\mathbb{R}^n$ and $\rho \geq \bar{\rho}$.
An observation (7.56) concerning the metric projection of a point to different sets has been made in [33], namely

$$\|pr_C y - pr_D y\| \leq 5\rho^{1/2} \hat{d}_\rho^{1/2}(C, D)$$

where $\rho \geq \max\{1/6, \|y\| + d(y, C) + d(y, D)\}$. Then we have

$$\begin{aligned} \|f(x) - f(\bar{x})\| &= \|pr_{F(x)} y(x) - pr_{F(\bar{x})} y(\bar{x})\| \\ &\leq \|pr_{F(x)} y(x) - pr_{F(x)} y(\bar{x})\| + \|pr_{F(x)} y(\bar{x}) - pr_{F(\bar{x})} y(\bar{x})\| \\ &\leq \|y(x) - y(\bar{x})\| + 5\rho^{1/2} \hat{d}_\rho^{1/2}(F(x), F(\bar{x})) \end{aligned}$$

Theorem 4.34 from the same monograph states that $\hat{d}_\rho^{1/2}(F(x), F(\bar{x})) \underset{x \to \bar{x}}{\longrightarrow} 0$ if and only if $F(x) \underset{x \to \bar{x}}{\longrightarrow} F(\bar{x})$ in the sense of Kuratowski-Painlevé. By assumption, this is satisfied. In this case $d(y(\bar{x}), F(x)$ converges to $d(y(\bar{x}), F(\bar{x})$ and, therefore, for x close to $\bar{x}$ the chosen $\bar{\rho}$ is large enough. Thus, applying the cited theorem and the continuity of y we get the continuity of f. □

Theorem 4.3 *Let $F : X \rightrightarrows \mathbb{R}^n$ be lower semi-continuous multifunction with non-empty closed convex images and closed graph. We assume that for all $x \in X$ there is a $y \in \mathbb{R}^n$ such that F is semi-differentiable at (x, y) and the semi-derivative $DF(x, y; h) \neq \{0\}$ for any direction $h \in X$.*
Given such a point $(\bar{x}, \bar{y})$ there exists a continuous selection f of F with $f(\bar{x}) = \bar{y}$, which is Hadamard-directionally differentiable at $\bar{x}$ with a derivative

$$f'(\bar{x}; h) \in DF(\bar{x}, \bar{y}; h), \ \forall h \in X.$$

Proof: Let $(\bar{x}, \bar{y}) \in graphF$ be the chosen point. We define a neighbourhood $U_{\bar{x}}$ of $\bar{x}$ by fixing some $\varepsilon > 0$ and setting $U_{\bar{x}} = \bar{x} + \varepsilon B$. Here B is the open unit ball in X. $DF(\bar{x}, \bar{y}; \cdot)$ is a continuous set-valued mapping , by Lemma 4.1 Hence, there exists a continuous selection $\bar{z}(h) \in DF(\bar{x}, \bar{y}; h)$. Let us consider the restriction of $\bar{z}(\cdot)$ on the unit sphere and define $z(\cdot)$ by setting:

$$\begin{aligned} z(h) &= \bar{z}(h) \text{ for } h : \ \|h\| = 1, \\ z(h) &= \|h\| \bar{z}(h/\|h\|) \text{ for } h : \ 0 < \|h\| < 1, \\ z(0) &= 0 \end{aligned}$$

Consequently, $z(\cdot)$ is a continuous positive homogeneous mapping, defined on the closed unit ball of X. It is obvious that $z(h) \in DF(\bar{x}, \bar{y}; h)$.
Pick up a point $x \in U_{\bar{x}}$. It can be represented as

$$x = \bar{x} + \|x - \bar{x}\| \frac{x - \bar{x}}{\|x - \bar{x}\|} = \bar{x} + \lambda h$$

by setting

$$\lambda = \|x - \bar{x}\| \text{ and } h = \frac{x - \bar{x}}{\|x - \bar{x}\|}$$

Consider $F(x), \ x \in U_{\bar{x}}$ and choose $y \in F(x)$ that are metric projections of $\bar{y} + z(x - \bar{x})$ onto $F(x)$. We define

$$f_{\bar{x}}(x) = pr_{F(x)}[\bar{y} + z(x - \bar{x})].$$

Having in mind Lemma 4.2, we have to clarify that for any point $\tilde{x} \in U_{\bar{x}}$ one has $F(x) \underset{x \to \tilde{x}}{\longrightarrow} F(\tilde{x})$ in order to prove the continuity of $f_{\bar{x}}$ at $\tilde{x}$. By the lower semi-continuity of F, we have

$$F(\tilde{x}) \subseteq \liminf_{x \to \tilde{x}} F(x) \text{ (cf. Lemma 6.2.2. in [8])}$$

On the other hand,

$$\limsup_{x \to \tilde{x}} F(x) \subseteq F(\tilde{x})$$

since F has a closed graph. Consequently,

$$F(\tilde{x}) = \limsup_{x \to \tilde{x}} F(x) = \liminf_{x \to \tilde{x}} F(x)$$

and $f_{\bar{x}}$ is continuous at any point of $U_{\bar{x}}$.
We consider further all points $x \in X \ : \ \|x - \bar{x}\| \geq \varepsilon$. There exists a $y \in F(x)$: F is

semi-differentiable at (x, y) and we can construct a neighbourhood U_x and a continuous mapping $f_x : U_x \to I\!R^n$ in the same way. We set $U_x^c = X \backslash U_{\bar{x}}$ and define neighbourhoods

$$\begin{aligned} V_{\bar{x}} &= U_{\bar{x}}, \\ V_x &= \frac{\varepsilon}{2} U_x \text{ for all } \ x \in U_x^c. \end{aligned}$$

It is clear that

$$(\bigcup_{x \in U_x^c} V_x) \cup V_{\bar{x}} = X.$$

Let $\{g_\alpha\}_\Lambda$ be the partition of the unity subordinated to this covering. This means that any g_α corresponds to a neighbourhood V_x. We denote this neighbourhood by V_α and the mapping f_x by f_α. Now set:

$$f(x) = \sum_{\alpha \in \Lambda} g_\alpha(x) f_\alpha(x).$$

$f(\cdot)$ is well-defined by $supp g_\alpha \subseteq V_\alpha$, where f_α is defined. Note also that f is continuous as a sum of continuous functions. Given a point x, $\sum_{\alpha \in \Lambda} g_\alpha(x) f_\alpha(x)$ is a finite sum by the local finiteness of the partition of unity. It is a convex combination of points $f_\alpha(x) \in F(x)$ and, thus, $f(x) \in F(x)$, i.e f is a continuous selection of F.
Since the point $\bar{x}$ and some part of $V_{\bar{x}}$ are not overlapped by other open sets V_x, there exists a function $g_{\bar{\alpha}}$ such that $g_{\bar{\alpha}}(x) = 1$ for all x in some smaller neighbourhood $\bar{V} \subseteq V_{\bar{x}}$. Consequently, $f(x) = f_{\bar{x}}(x)$ on $\bar{V}$. Therefore, $f(\bar{x}) = \bar{y}$ and f is Hadamard-directionally differentiable at $\bar{x}$ with $f'(\bar{x}; h) \in DF(\bar{x}, \bar{y}; h), \ \forall h \in X$, by the construction of $f_{\bar{x}}$.
This completes the proof. □

Now some comparison between the two results is necessary. An advantage of the last construction is that the assumptions imposed on the set-valued mapping are milder than those required in the Theorem 3.11. The disadvantages are

- the implicit character of the construction - there is no formula for computing the value of the selection and its directional derivative;
- the constructed selection is directionally differentiable only at a chosen point, the differentiability property cannot be globalized. The only possible globalization is the following:

Remark 4.4 *Under the assumptions of the previous theorem suppose that $x_i \in X, \ i \in I$, where I is some index-set, are such that*

(i) there is $\varepsilon > 0$ with $B(x_i, \varepsilon) \cap B(x_j, \varepsilon)$ $\forall i \not\parallel$,

(ii) $\forall i \in I$ there is a $y_i \in F(x_i)$: F is semi-differentiable at (x_i, y_i) with a semi-derivative $DF(x_i, y_i; h) \neq \{0\}$ for any $h \in X$.

Then there is a continuous selection f of F such that $\forall i \in I \ f(x_i) = y_i$ and f is Hadamard-directionally differentiable at x_i with $f'(x_i; h) \in DF(x_i, y_i; h), \ \forall h \in X$.

In [15], Dommisch establishes two results on the topic. The existence of a smooth partition of unity in $\mathbb{R}^n$ is exploited there to prove the existence of a smooth selection in the following sitations:

- F acts from a closed subset of $\mathbb{R}^m$ in $\mathbb{R}^n$ and has convex images and open preimages, i.e., $F^{-1}(y) = \{x \in X : y \in F(x)\}$ is open in X for all $y \in \mathbb{R}^n$.
- The graph of F has a smooth boundary in the sense that the support function is continuously differentiable everywhere.

Applications of these results for mappings arising from optimization problems are difficult. The assumptions in the first statement imply the openness of the graph of the multifunction. The graph of a set-valued mapping is closed under relatively mild assumptions and it is not natural to require some stronger regularity like differentiability, supposing the graph to be open. The second statement is obtained by using a selection similar to the Steiner selection and therefore it is not a surprising result, but it assumes conditions, that are hard to verify and even to fulfil. Moreover, we would like to keep X a (possibly) infinite dimensional space, which is essential for the application we present in the section 6.

5 Measurability, Castaing representation and Delta-Theorems

We equip the linear normed space X with a σ-algebra $\mathcal{A}$. In our further considerations this will be mainly the Borel σ-algebra, which we denote by $\mathcal{B}(X)$. We use the following definition of measurability (see also [8, 6]):

Definition 5.1 *A mapping $f : (X, \mathcal{A}) \to \mathbb{R}^n$ is measurable if for any open set $C \subseteq \mathbb{R}^n$ the preimage $f^{-1}(C) = \{x \in X \ : \ F(x) \in C\}$ belongs to $\mathcal{A}$.*
A multifunction $F : (X, \mathcal{A}) \rightrightarrows \mathbb{R}^n$ is measurable if for any open set $C \subseteq \mathbb{R}^n$ $F^{-1}(C) = \{x \in X \ : \ F(x) \cap C \neq \emptyset\} \in \mathcal{A}$.

Recall that $f : (X, \mathcal{A}) \to \mathbb{R}^n$ is called *a measurable selection* of F if f is measurable and $f(x) \in F(x)$ almost surely.
It is known (cf. [23]), that a closed-valued measurable multifunction in a Polish target space possesses a measurable selection. Furthermore, for a multifunction F with nonempty closed values in a Polish target space (in our case this is $\mathbb{R}^n$), we can choose a countable family of measurable selections $\{f_n\}$ that pointwise fills up the values of the multifunction:

$$\text{for each } x \in X, \ F(x) = cl(\cup_{n=1}^{\infty} f_n(x)).$$

Such a countable family is called *a Castaing representation* for the multifunction. Moreover, the existence of such a representation characterizes measurability.

Our aim is to construct measurable selections, respectively Castaing representations of a multifunction $F : X \rightrightarrows \mathbb{R}^n$ with additional regularity properties, namely directional differentiability. Why this is of interest will become clear when we take into consideration the so-called delta-approach. Delta-theorems are concerned with the asymptotic distribution of functions of random elements.

Theorem 5.2 *Let $f : (X, \mathcal{B}(X)) \to \mathbb{R}^n$ be measurable and Hadamard-directionally differentiable at some point $\bar{x} \in X$. Suppose that $\tau_n(x_n - \bar{x})$ are some random elements of X converging in distribution to some element z, written*

$$\tau_n^{-1}(x_n - \bar{x}) \xrightarrow{d} z$$

while $\tau_n \downarrow 0$ and z is a random element in some separable subspace of X. Then

$$\tau_n(f(x_n) - f(\bar{x})) \xrightarrow{d} f'(\bar{x}; z).$$

Here $\xrightarrow{d}$ denotes convergence in distribution.
Recall that convergence in distribution of a sequence of random elements x_n,
$x_n : (\Omega, \mathcal{A}, P) \to X$ means the weak* convergence of the measures $\mu_n = P \circ x_n^{-1}$ these elements unduce on the space X. A sequence of probability measures μ_n on a separabel metric space X weakly* converges to μ (cf.[9]) if

$$\lim_{n\to\infty} \int g(x)\mu_n(dx) = \int g(x)\mu(dx)$$

for all bounded continuous functions $g : X \to \mathbb{R}$. Convergence in distribution of set-valued mappings is considered in [35] and we recall some facts for the sake of completeness and for the convenience of the reader.
A closed-valued multifunction $F : (\Omega, \mathcal{A}) \rightrightarrows \mathbb{R}^n$ can be identified with a single valued mapping $\tilde{F} : (\Omega, \mathcal{A}) \to \mathcal{F}(\mathbb{R}^n)$, where $\mathcal{F}(\mathbb{R}^n)$ is the hyperspace of all closed subsets of $\mathbb{R}^n$ equipped with the topology of Kuratovski-Painlevé-convergence. It is known that $\mathcal{F}(\mathbb{R}^n)$ is then a Polish space. Measurability of F is equivalent to measurability of $\tilde{F}$. Then, in a natural way, convergence in distribution of F_n to some F is understood as the weak* convergence of the measures unduced on $\mathcal{F}(\mathbb{R}^n)$ by $\tilde{F}_n$. Relying on the fact that a sequence of closed sets converges in $\mathcal{F}(\mathbb{R}^n)$ if and only if the sequence of the distance functions converges pointwise, Salinetti and Wets demonstrated in [35], Theorem 25 that a sequence of multivalued mappings $\{F_n\}$ converges in distribution if and only if the distance processes $\{d(\cdot, F_n)\}$ converge as stochastic processes on $\mathbb{R}^n$. The first generalized delta-theorem for set-valued mappings was formulated by King [21]. Other versions are given in [38] by Shapiro and in [24] by Lachout.

Theorem 5.3 *Let $F : (X, \mathcal{B}(X)) \rightrightarrows \mathbb{R}^n$ be closed-valued measurable multifunction defined on a separable complete metric space X. Suppose that x_n satisfy a generalized central limit formula with limit $\bar{x}$, i.e., there is a sequence $\{t_n\}, t_n \geq 0$ monotonically decreasing to 0 and a limit distribution z such that*

$$t_n^{-1}(x_n - \bar{x}) \xrightarrow{d} z$$

as random variables in X.
Assume, additionally, that F is almost surely semi-differentiable at $(\bar{x}, \bar{y})$ for some $\bar{y} \in F(\bar{x})$ with respect to the measure induced by z. Then $\{F(x_n)\}$ satisfies the generalized central limit formula

$$t_n^{-1}(F(x_n) - \bar{y}) \xrightarrow{d} DF(\bar{x}, \bar{y}; z)$$

as random closed sets in $\mathbb{R}^n$ or, equivalently,

$$d(\cdot, t_n^{-1}[F(x_n) - \bar{y}]) \xrightarrow{d} d(\cdot, DF(\bar{x}, \bar{y}; z))$$

as stochastic processes on $\mathbb{R}^n$.

Here semi-differentiability almost surely means that the convergence of the differential quotients holds for all but a set of measure 0 directions.
Some natural question arise: What is the behaviour of the measurable selections of F if F satisfies the central limit formula above? Do these selections satisfy a similar central limit formula? Do the delta-theorem for single-valued mappings apply to them?
A general answer to these questions is not known. The main results of this section will contribute to the investigations of this matter. First, we will formulate theorems on Castaing representations with differentiability properties. We will see that the two different constructions lead to substantially different statements.

Theorem 5.4 *Let X be a linear normed space and $F : X \rightrightarrows \mathbb{R}^n$ be a measurable multifunction with non-empty compact convex images. Suppose that F is locally Lipschitzian at any point $x \in X$ and semi-differentiable at any point $(x, y) \in graphF : \ y \in bdF(x)$. Then F admits a Castaing representation by (Steiner) selections $\{f_n\}$ with the following property:*
For any point $(x, y) \in graphF$ and for any $\varepsilon > 0$ there is a selection f_n such that $\|y - f_n(x)\| \leq \varepsilon$ and f_n is Hadamard-directionally differentiable at x with derivative satisfying the formula 3.
If F is semi-differentiable at $(x, f_n(x))$ then $f_n'(x; h) \in DF(x, f_n(x); h), \ \forall h \in X$.

Proof: Let $\{y_n\}$ be a dense set of points in $F(X)$. Now, pick up a point y_n. There exists an $x_n : y_n \in F(x_n)$. A probability measure α_n on Σ can be chosen such that $y_n = st(F(x_n))$ using the measure in the defining formula for the Steiner center. By Theorem 3.11 the Steiner selection (defined with respect to the measure α_n) is Hadamard-directionally differentiable and has a value y_n. Moreover, the Steiner selection is measurable (see [6]). We denote it by f_n. Evidently, the family $\{f_n\}$ of selections constructed through the points y_n is a Castaing representation of F with the desired property. □

Theorem 5.5 *Let X be a separable normed linear space and $F : X \rightrightarrows \mathbb{R}^n$ be a measurable lower semi-continuous multifunction with non-empty closed convex images and closed graph. Suppose that F is semi-differentiable at all points $(x, y) \in graphF$. Then F admits a Castaing representation by continuous selections $\{f_n\}$ with the following property:*

For any point $(x, y) \in graphF$ and for any $\varepsilon > 0$ there is a selection f_n such that $\|y - f_n(x)\| \leq \varepsilon$ and f_n is Hadamard-directionally differentiable at some $\tilde{x} : \|x - \tilde{x}\| \leq \varepsilon$ with $f_n'(\tilde{x}; h) \in DF(\tilde{x}, f_n(\tilde{x}); h),\ \forall h \in X$.

Proof: By the separability of X, there exist countably many points x_n which span a dense subset of X. Consider $F(x_n)$ and choose $\{y_n^i\}$ to be countable set of points that are dense in $F(x_n)$. Fixing a point (x_n, y_n^i) we apply Theorem 4.3 to construct a continuous selection f_{ni} through $\{y_n^i\}$ and Hadamard-directionally differentiable at x_n with directional derivatives belonging to the corresponding semi-derivative of F. The selection f_{ni} is measurable by its continuity. The family $\{f_{ni}\}$ is the Castaing representation of F we are looking for. □

Note that the semi-differentiability of F is actually not needed at all points of its graph.It is enough to have this property at a set of points that is countable and dense in $graphF$.

The first theorem imposes stronger restrictions on the multifunction under consideration. On the other hand, in this case we get a family of selections that are directionally differentiable everywhere due to the properties of the Steiner selection. In the second theorem, we get a family of selections that provide directional differentiability at some points, which are close to the points we consider. However, if we are interested in obtaining a Castaing representation with differentiability properties at some particular point, the construction of the latter theorem works:

Corollary 5.6 *Suppose the assumptions of Theorem 5.5 and let $x \in X$ be a given point. Then F admits a Castaing representation $\{f_n\}$ with the properties stated there and, moreover, all of the functions f_n are Hadamard-directionally differentiable at x.*

Results of this kind are interesting for their application to obtain central limit formulas for selections of multifunctions. Now, we are ready to come back to this problem.

Proposition 5.7 *Let $F : X \rightrightarrows \mathbb{R}^n$ be a measurable multifunction and suppose that either the assumptions of Theorem 5.4 or those of Theorem 5.5 hold true. Let us assume that the random elements $x_n \in X$ satisfy a generalized central limit formula with limit $\bar{x}$, i.e.,*

$$t_n^{-1}(x_n - \bar{x}) \xrightarrow{d} z$$

as random variables in X.
Given a point $\bar{y} \in F(\bar{x})$, then $\{F(x_n)\}$ satisfies the generalized central limit formula

$$t_n^{-1}(F(x_n) - \bar{y}) \xrightarrow{d} DF(\bar{x}, \bar{y}; z)$$

as random closed sets in $\mathbb{R}^n$.
Moreover, F admits a Castaing representation $\{f_k\}$ such that all f_k satisfy the generalized central limit formula

$$t_n^{-1}[f_k(x_n) - f_k(\bar{x})] \xrightarrow{d} f_k'(\bar{x}; z) \in DF(\bar{x}, f_k(\bar{x}); z)$$

as random variables on $\mathbb{R}^n$.

Proof: The proof follows from the previous corollary and the Theorem 5.3 and Theorem 5.4. □

Let us mention some of the results published on the topic convergence in distribution of measurable selections of multifunctions. Interesting results are given in [1] by Artstein, where the set of distributions of measurable selections of random sets has been studied from a slightly different point of view. Let us first emphasize that the distribution of a random set does not determine the ensemble of distributions of its measurable selections. This phenomenon is discussed in [1] and one can also find more references on the matter cited therein. The primal object in [1] is a given distribution on some compact subset of a complete separable metric space. The question is considered which distributions on the space are induced by selections of random sets with the given probability distribution. It has been shown that these distributions build a convex compact set, which depends continuously on the given distribution. The conclusion for our research on proving central limit formulas for selections (like in the previous proposition) is that we cannot expect to get some formula for all the measurable selections of the mapping F. In [2] "almost" Castaing representations are considered.

Comparable results are given by King [21] and Lachout [24]. In Theorem 4.3 in [21] a generalized central limit formula for all measurable selections is established under the assumption that the multifunction has a kind of Lipschitz behaviour, $F(\bar{x}) = \bar{y}$, and $DF(\bar{x}, \bar{y}; z)$ is single-valued almost everywhere. As noticed in the same paper, under these conditions F is not far from being single-valued in a small neighbourhood of $\bar{x}$. In our paper we would like to keep the multivaluedness of the mapping under consideration. This is essential for the application of our result (Proposition 5.7) we will present in the next section.

Another relevant result is contained in the paper [24]. A multifunction acting between two infinite-dimensional spaces is considered there. The values $F(x)$ are supposed to be compact and $F(\bar{x}) = \bar{y}$ to be a singleton. The statement is that the measurable selections f of F do not satisfy the central limit formula themselves, but there are subsequences for which the formula holds.

Summarizing we can say that we do not have a result stating a central limit formula for all selections of the multifunction, but we are able to construct a Castaing representation with this property under relatively general assumptions. Having in mind the literature mentioned above, it is clear that statements of this kind should be possible for all measurable selections under very restrictive assumptions only.

6 Two-stage stochastic programs

In this section we present a non-trivial example of a set-valued mapping, which satisfies the assumptions of the main results of the previous sections. We will consider stochastic optimization problems subjected to perturbations. The solution set mapping of these problems will be the subject of our investigations. We shall see how Proposition 5.7 leads to statements about the asymptotic behaviour of solution sets.

Let us consider a two-stage stochastic program with linear recourse and random right-

hand side:

$$\min\{g(x) + Q_\mu(Ax) : x \in C\} \tag{7}$$

$$Q_\mu(\chi) = \int_{\mathbb{R}^m} \tilde{Q}(\theta - \chi)\mu(d\theta), \tag{8}$$

$$\tilde{Q}(z) = \min\{q^\top y : Wy = z, y \geq 0\} \tag{9}$$

where $g : \mathbb{R}^n \to \mathbb{R}$ is a convex function, $C \subseteq \mathbb{R}^n$ is a non-empty closed convex set and μ is a Borel probability measure on $\mathbb{R}^m$. Furthermore, $q \in \mathbb{R}^s$ and $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, $W \in L(\mathbb{R}^s, \mathbb{R}^m)$. We make use of the general assumptions (A1)-(A3), which are common in the literature in order to make the problem well-defined.

(A1) $W(\mathbb{R}^s_+) = \mathbb{R}^m$ (complete recourse),

(A2) $M_D := \{u \in \mathbb{R}^m : W^\top u \leq q\} \neq \emptyset$ (dual feasibility),

(A3) $\int_{\mathbb{R}^m} \|z\|\mu(\mathrm{d}z) < +\infty$ (finite first moment).

Having in mind linear programming theory, observe that (A1) and (A2) imply $\tilde{Q}(z)$ to be finite for all $z \in \mathbb{R}^m$. Due to (A3) also the integral of $\tilde{Q}(z)$ is finite ([20, 42]).

The model is derived from an optimization problem with uncertain data, where some statistical information about the random data is available. The decision x of the first stage has to be made *here and now* before observing some realization of θ. It is supposed to solve the problem:

$$\inf\{g(x) : x \in C, Ax = \theta\}$$

After observing a realization of θ we fix a second-stage decision y (called recourse action) in order to overcome the deviation $\theta - Ax$. The matrix W determines the rule to react and q the costs of our reaction. (A1) means that we are able to overcome any deviation. To choose y properly, we minimize its costs. To choose x properly, we minimizes the sum of the first-stage costs and the expected second-stage costs, caused by the corrective action y. Further details and fundamental properties of (two-stage) stochastic programs can be found in [20, 42].

Precise information about the measure μ in practical applications is difficult to be gained. This is the reason for the stability analysis of the problem if the measure μ (and hence Q_μ) is subjected to perturbations. The following setting of perturbed problems will be considered:

$$\inf\{g(x) + Q(Ax) : x \in C\},$$

where $Q : \mathbb{R}^m \to \mathbb{R}^n$ is a convex function, considered to be a perturbation (resp. approximation) of the expected recourse function Q_μ. Resorting to convex perturbations is motivated by the fact that, given (A1) and (A2), Q_μ is convex for any probability measure with finite first moment (cf. [20, 42]).

Let us consider the solution sets of the perturbed problems. We are interested to have

some selections of them such that gaining information about the asymptotic behaviour of the selections at the point Q_μ is possible. We denote the solution set mapping by

$$\psi(Q) = \operatorname{argmin}\{g(x) + Q(Ax) : \ x \in C\}$$

Two-stage stochastic programs hardly have a unique solution. This fact has motivated the attempt to avoid the assumption on the multifunction to be a singleton at certain points in our investigations. The next example give an impression on how restrictive this assumption is.

Example 6.1 $g(x) = 0,\ A = (1,0),\ C = [0,1] \times [0,1],\ q = (1,1),\ W = (1,-1)$.
Let μ be the uniform distribution on [-1/2,1/2]. Then

$$\begin{aligned}
&\psi(Q_\mu) \\
&= \operatorname{argmin}\{Q_\mu(Ax) : \ x \in [0,1] \times [0,1]\} \\
&= \operatorname{argmin}\{\int_{I\!R} |\omega - x_1| \mu(d\omega) : \ x \in [0,1] \times [0,1]\} \\
&= \{(0, x_2) : x_2 \in [0,1]\} = ker A \cap C
\end{aligned}$$

One can see that even for very simple examples the solution set is not a singleton. Under an assumption that Q_μ is a strictly, respectively strongly convex function we have the uniqueness of $A\psi(Q_\mu)$, but we cannot expect that $ker A = \{0\}$.

Supposed we have some approximations (resp. estimates) $Q_n,\ n = 1, 2, \ldots$ of Q_μ that satisfy a generalized central limit formula in some functional space. In order to apply the delta-theorem approach we have to show first that (under some conditions) the solution set mapping is semi-differentiable at (Q_μ, x) for some (or for all) $x \in \psi(Q_\mu)$. This question is studied in [13], where sufficient conditions for the semi-differentiability of ψ at all points $(Q_\mu, x) : \ x \in \psi(Q_\mu)$ are given.

Setting $D = A(C)$, we consider two functional spaces. The space $C^{1,1}(D, R)$ of all realvalued continuously differentiable functions with locally Lipschitz-continuous derivative, defined on D, and the space $C^{0,1}(D, R)$ of all real valued locally Lipschitz-continuous functions, defined on D. Both spaces are metrizable (cf. [13]). We suppose here that the set C is bounded and endow the space $C^{0,1}(D, R)$ with the usual Lipschitz-norm:

$$\|f\| = \|f\|_\infty + \sup_{x,y \in D} \frac{|f(x) - f(y)|}{\|x - y\|}$$

We work with the corresponding norm-convergence in $C^{1,1}(D, R)$.

In the following, we always consider the restriction of the solution set mapping ψ to the cone K of convex functions in one of the spaces above. This has an impact on the notion of semi-differantiability. Considering the semi-derivative at a point (Q_μ, x) in a certain direction $\bar{v}$, we assume that the arguments of ψ lie in K. Hence, $Q_\mu + tv \in K\ \forall v \to \bar{v}$. Consequently, v are elements of the closure of the radial tangent cone to K at the point Q_μ. We denote the radial tangent cone to K at the point Q_μ by

$$T_K^r = \{\lambda(Q - Q_\mu) : \ \lambda \geq 0,\ Q : I\!R^n \to I\!R \text{ convex}\}.$$

Theorem 6.2 *[13]: Assume $\psi(Q_\mu)$ to be non-empty and bounded. Let Q_μ be strictly convex on some open neighbourhood of $A(\psi(Q_\mu))$ and twice continuously differentiable at χ_* : $A(\psi(Q_\mu)) = \{\chi_*\}$. Let g be twice continuously differentiable, $v \in T_K^r(Q_\mu)$ and $x \in \psi(Q_\mu)$.*
Assume additionally that:

(i) there exists $\delta > 0$ such that for all $t \in [0, \delta)$
$\mathrm{d}(x, \psi(Q_\mu + tv)) = O(t)$,

(ii) there exists a neighbourhood U of $\psi(Q_\mu)$ and constants $c > 0,\ \delta > 0$ such that:

$$g(x) + (Q_\mu + tv)(Ax) \geq \varphi(Q_\mu + tv) + c\mathrm{d}(x, \psi(Q_\mu + tv))^2, \quad \forall x \in U \cap C,\ \forall t \in [0, \delta].$$

(iii) for all $y \in S(x)$, where

$$S(x) = \{y \in T_C(x) : \nabla g(x)y + \nabla Q_\mu(Ax)Ay = 0\}$$

the corresponding second-order set $S^2(x, y)$ is non-empty,

$$S^2(x, y) = \{z \in T_C^2(x, y) : \nabla g(x)z + \nabla Q_\mu(Ax)Az = 0\}$$

Then ψ is directionally differentiable at Q_μ in direction v and

$$\begin{aligned} \psi'(Q_\mu; x)(v) &= \lim_{t \to 0+} \frac{1}{t}(\psi(Q_\mu + tv) - x) \\ &= \operatorname{argmin}\left\{\frac{1}{2}\langle \nabla^2 g(x)y, y\rangle + \frac{1}{2}\langle \nabla^2 Q_\mu(Ax)Ay, Ay\rangle + v'(Ax; Ay) : y \in S(x)\right\}. \end{aligned}$$

The consideration at this place is in the space $C^{0,1}(D, R)$.
The condition (i) requires a kind of Lipschitz behaviour for the solution set-mapping, (ii) is a second-order growth-condition, and (iii) imposes a restriction on the second order tangent set to C at the optimal points with respect to some elements of the tangent cone. These conditions are verified in [13] for the particular case where g is, in addition, a quadratic function and C is a polyhedral set. In this case also a strengthened version of condition (i) is satisfied, namely:

(i)$^+$ there exists $\delta > 0$ such that for all $t \in [0, \delta)$
$d_H(\psi(Q_\mu), \psi(Q_\mu + tv)) = O(t)$

Let $C^{0,1}(D, R)$ be equipped with its Borel σ-algebra.

Proposition 6.3 *Under the conditions of the previous theorem, assuming (i)$^+$ instead of (i), the mapping ψ possesses a Castaing representation $\{f_n\}$ of Steiner selections that are directionally differentiable at Q_μ in directions $v \in T_K^r(Q_\mu)$.*

Proof: We shall have established the proposition, if we verify the conditions of Corollary 3.13. Then the proof of Theorem 5.4 can be carried over. First, we observe that $\psi(Q)$ are compact convex sets for all Q of the domain of ψ by the properties of convex programs and by the assumption that C is bounded. Hence, Steiner points of $\psi(Q)$ are well-defined. Now, we pick up points x_n, which build a countable dense set in the image of ψ. Let Q_n denote those points for which $x_n \in \psi(Q_n)$. Analogously to Theorem 5.4, we construct Steiner selections f_n through x_n that form a Castaing representation. Now we are left with the task to prove that f_n are directionally differentiable at Q_μ in directions $v \in T_K^r(Q_\mu)$. This property follows from the Corollary 3.13, taking into account that the proof of the directional differentiability of the Steiner selection in some direction requires the directional differentiability of the multifunction in this direction, only. □

Now we come to the semi-differentiability of the solution set-mapping and its consequences. From now on, we consider the restriction of ψ on the space $C^{1,1}(D,R)$.

Theorem 6.4 *[13]: Assume $\psi(Q_\mu)$ to be non-empty, g be a quadratic function and C be a polyhedron. Let Q_μ be strongly convex on some open neighbourhood of $A(\psi(Q_\mu))$ and twice continuously differentiable at χ_* : $A(\psi(Q_\mu)) = \{\chi_*\}$. Let $x \in \psi(Q_\mu)$.*
Then ψ is semi-differentiable at (Q_μ, x) and the semi-derivative is given by the formula of the previous theorem.

This result has the following consequences with respect to the delta-approach:

Theorem 6.5 *Under the conditions of the previous theorem suppose that Q_n, $n = 1, 2, \ldots$ satisfy the functional central limit formula*

$$t_n^{-1}[Q_n - Q_\mu] \xrightarrow{d} \zeta \quad in \quad C^{1,1}(D,R)$$

for some monotonically decreasing sequence $t_n \downarrow 0$. Given a point $\bar{x} \in \psi(Q_\mu)$, then ψ satisfy the generalized central limit formula

$$t_n^{-1}[\psi(Q_n) - \bar{x}] \xrightarrow{d} D\psi(Q_\mu, \bar{x}; \zeta)$$

as random sets in $\mathcal{F}(I\!R^n)$.
Moreover, ψ admits a Castaing representation $\{f_k\}$ of Steiner selections such that all f_k satisfy the central limit formula

$$t_n^{-1}[f_k(Q_n) - f_k(Q_\mu)] \xrightarrow{d} f_k'(Q_\mu; \zeta) \in D\psi(Q_\mu, f_k(Q_\mu); \zeta)$$

as random variables on $I\!R^n$ and there is a k_0 such that $f_{k_0}(Q_\mu) = \bar{x}$.

Proof: The mapping ψ is semi-differentiable at Q_μ by virtue of the previous theorem. The assertion follows from Proposition 5.7 applying Corollary 3.12. □

Investigating the asymptotic behaviour of solution sets of stochastic programs is beyond the scope of this paper. The last statements have been included for the sake of giving an application of the results of this paper and yielding non-trivial statements.

For investigations on the asymptotic behaviour of stochastic programs the interested reader is referred to [16, 27, 22, 38] and the references therein.

Acknowledgment: The author wishes to express her gratitute to W.Römisch for many stimulating conversations and his interest while this work was in progress.

References

[1] Z. Artstein, Distributions of random sets and random selections, *Israel J. Math.* 46 (1984) 313-324

[2] Z. Artstein, On dense univalued representations of mulivalued maps, *Rendiconti der Circolo Mathematico di Palermo*, Serie II, Tomo XXXIII, (1983) 340-350

[3] J.-P. Aubin, Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusions and differential inclusions, Math. Analysis and Applications, Part A: Advances in Math. Supplementary Studies, Vol.7A, Academic Press, 1981,160-232.

[4] J.-P. Aubin, Lipschitz behaviour of solutions to convex minimization problems, *Mathematics of Operations Research* 8 (1984) 87-111.

[5] J.-P. Aubin and A.Cellina, Differential Inclusions, Springer, Grundlehren der math. Wiss., 1984, Vol.264.

[6] J.-P. Aubin and H. Frankowska, Set-Valued Analysis, Birkhäuser, Boston 1990.

[7] A. Auslender and R. Cominetti, A comparative study of multifunction differentiability with applications in mathematical programming, *Mathematics of Operations Research*, 16 (1991), 240-258.

[8] G. Beer, Topologies on Closed and Closed Convex Sets, Kluwer, Mathematics and its Applications, Vol.268, 1993

[9] P.Billingsley, Convergence of probability measures, Wiley, New York, 1968.

[10] Ch. Castaing, Sur les multi-application measurables, *Inf. Rech. Op.* 1 (1967) 91-126

[11] Ch. Castaing and M. Valadier Convex analysis and measurable multifunctions, Springer, Lecture Notes in Math., 1977, Vol.580.

[12] R. Cominetti, Metric regularity, tangent sets and second-order optimality conditions, *Applied Mathematics and Optimization* 21 (1990), 265-287.

[13] D.Dentcheva and W.Römisch, Differential stability of two-stage stochastic programs, manuscript, Humboldt University, Berlin, 1996

[14] F.Deutsch and P. Kenderov, Continuous selections and approximate selections for set-valued mappings and applications to metric projections, *SIAM J. Math. Anal.* 14 (1983), 183-194

[15] G.Dommisch, On the existence of Lipschitz-continuous and differentiable selections for multifunctions, *in* J.Guddat et al.(eds.): Parametric Optimization and related topics, Math. Res. Vol. 35, Akademie-Verlag Berlin, 1987, 60 - 73.

[16] J. Dupačová and R.J.-B. Wets, Asymptotic behaviour of statistical estimators and of optimal solutions of stochastic optimization problems, *Annals of Statistics* 1988, 1517-1549.

[17] H. Frankowska, Inclusions adjointes associeex aux trajectoires minimales d'inclusions differentielles, *C.R.Acad. Sc. Paris* 297 (1983) 461-464

[18] H. Frankowska, Local controllability and infinitesimal generators of semigroups of set-valued maps, *SIAM J. Control Opt.* 25 (1987), 412-432.

[19] S. Gautier and R. Morchadi, A selection of convex-compact-valued multifunctions with remarkable properties: the Steiner selection, *Numer. Funct. Anal. and Optimiz.* 13(5& 6) (1992) 513-522

[20] P. Kall, Stochastic Linear Programming, Springer-Verlag, Berlin 1976.

[21] A.J.King, Generalized delta theorems for multivalued mappings and measurable selections, *Mathematics of Operations Research* 14 (1989) 4, 720-736

[22] A.J.King and R.T.Rockafellar, Asymptotic theory for solutions in statistical estimation and stochastic programming, *Mathematics of Operations Research* 18 (1993) 1, 148-162.

[23] K. Kuratowski and C. Ryll-Nardzewski, A general theorem on selectors, *Bull. Acad. Pol. Sc.* 13, 1965, 397-403.

[24] P. Lachout, On multifunction transforms of probability measures, *Annals of Operations Research*, 56 (1995) 241-250.

[25] E. Michael, Selected selection theorems, *Amer. Math. Montly* 63 (1956) 233-237.

[26] J.-P. Penot, Differentiability of relations and differential stability of perturbed optimization problems, *SIAM J. Control and Optimization* 22 (1984) 4, 529-551

[27] G.Pflug, Asymptotic stochastic programs, *Mathematics of Operations Research*, 20 (1996) 4, 769-789

[28] R.A. Poliquin, Proto-differentiation of subgradient set-valued mappings, *Canad. J. Math.* 42 (1990) 520-532.

[29] R.A. Poliquin and R. T. Rockafellar, Amenable functions in optimization, *in:* F.Giannessi (ed.), Nonsmooth optimization methods and applications, Gordon & Breach, Philadelphia, 1992, 338-353.

[30] A. Prekopa Stochastic Programming, Kluwer, Math. and its Applications, Vol. 324, 1995

[31] R. T. Rockafellar, Proto-differentiability of set-valued mappings and its application in optimization, *in:* H. Attouch et al. (eds.), Analyse Non Lineaire, Gauthier-Villars, Paris, 1989, 449-482.

[32] R. T. Rockafellar, Convex Analysis, Princeton University Press, Princeton 1970.

[33] R. T. Rockafellar and R.J.-B.Wets, Variational analysis, forthcomming

[34] J. Saint-Pierre, Point de Steiner et selection Lipschitzienne, *Travaux de Seminaire d'analyse convexe - Montpellier* 185, Expose 7.

[35] G.Salinetti and R.J.-B.Wets, On the convergence in distribution of measurable multifunctions (random sets), normal integrands, stochastic processes and stochastic infima, *Mathematics of Operations Research* 11 (1986) 3, 385-419.

[36] A. Shapiro, Existence and differentiability of metric projections in Hilbert spaces, *SIAM J. Optimization* 4 (1994) 1, 130-141.

[37] A. Shapiro, On differentiability of metric projections in $\mathbb{R}^n$: Boundary Case, *Proc.Amer. Math. Sic.* 99, 123-128.

[38] A. Shapiro Asymptotic analysis of stochastic programs, *Annals of Operations Research* 30 (1991), 169-186.

[39] G.C.Shephard, A uniqueness theorem for the Steiner point of a convex region, *J. London Mat. Soc.* 43 (1968) 439-444.

[40] J.Steiner, Von der Krümmungsschwerpunkte ebener Kurven, *Zeitschrift für Reine Angew. Math.* 21 (1840) 33-63, 101-122.

[41] C. Ursescu, Tangent sets calculus and necessary conditions for extremality, *SIAM J. Control Opt.* 20 (1982), 563-574.

[42] R.J.-B. Wets, Stochastic programs with fixed recourse: the equivalent deterministic program, *SIAM Review* 16 (1974), 309-339.